

# Weekly Report

## 05/25/2015 - 05/31/2015

Jing XIA

May 31, 2015

## 1 Summary

This week I mainly focus on the DataScanner2 project (the second phase of data inspection).

## 2 Projects

### 2.1 Project 1 - DataScanner2

**DataScanner2:** A system that integrates various data quality metrics, covering all quality metrics categories and providing a quality-aware (clustering, correlation, outliers, complex patterns, image quality, feature preservation) visualization.

This week I mainly implemented the web interface:

**Bing map** For view or views of geo-locations, a map is necessary to link the data with. We implemented Google map in the last version but it no longer works due to Google service issue. Thus a Bing map is adopted in a div instead of an iframe as in the last version.

**Radar of metrics** Since multiple quality metrics are applied in view quality and view correlation evaluation, we want to visualize those metrics so that users can identify what are the main qualities for certain views.

**2D view logo** Originally all 2D views were represented with a scatter plot logo, which is weak in distinguishing 2D views of different combinations of variable types. We use the combination of one dimension logo to indicate the variable types of the underlying 2D views.

TO-DOs:

- a. Locate tree nodes with Voronoi treemap layout
- b. Filter views with the matrix widget

- c. Move the data table to the Vis tab
- d. Paper writing

More discussions:

- How to explore the layout for better revealing of redundant dimensions? The redundancy can be firstly inferred by locating neighboring views and then evaluating with their point-wise mutual information.
- How to explore the layout for better revealing conflict data? However, conflict data can only be uncovered with an integration of domain knowledge. How to introduce domain knowledge to the layout.

## **2.2 Project 2 - NBA Game Visualization**

-

## **2.3 Project 3 - FlowTracker**

We sent emails to the dataset provider for answers and got one reply but have not figured out the problem yet. We integrated the game data.

## **3 Paper Reading**

-

## **4 To Do List**

1. DataScanner2 — paper writing.
2. DataScanner2 — web interface implementation.
3. Group seminar report.